

Analysis and assessment are inherently judgmental processes that should be guided by the original evaluation objectives and professional standards commonly used by the evaluation and research community. An information source that might be useful to WIC administrators in choosing analytic methods is "The Evaluation Framework," a publication of the U.S. General Accounting Office.

Qualitative Data

Process evaluations (discussed earlier in section I) often involve the analysis of qualitative, descriptive data. Qualitative data includes descriptive information on such things as staff opinions on program redesign or the quality of services from a participant's perspective. The researcher collects data and looks for common themes, patterns, or hypotheses. In essence, qualitative analysis may not include counting but rather summarizing and interpreting study findings. Qualitative data can strengthen and further explain quantitative findings. Analyzing qualitative data usually requires substantial knowledge of the program as well as a thorough understanding of the tasks comprising day-to-day delivery of program services. An evaluator who assesses a large volume of qualitative data may wish to incorporate an advisory panel into the analysis portion of the evaluation who can provide a broad view and consensus.

Quantitative Data

More formally structured analysis methods are applied to quantitative or numerical data. These methods range from simple, easy-to-calculate descriptive statistics (discussed below) to sophisticated regression analyses such as probit and logit.

Measures of Central Tendency

There are various statistical techniques to analyze particular types of data. An evaluator must first determine which type of data is available for analysis. Usually, an evaluator who analyzes quantitative data first calculates *descriptive statistics*. Primary among such statistics are measures of central tendency which represent typical or "average" values of the variables under study. A measure of central tendency is valuable because it captures, in a single number, information on an entire population or sample. Measures of central tendency are defined below.

■ The mode is the value or item occurring most frequently in a series of observations or statistical data. The mode can be used with any type of data. If, for instance, WIC records indicate the distribution of WIC breastfed infants is 200 black infants, 400 white, and 100 Hispanic, then the modal, or "typical," breastfed infant is white.

■ The median is the value that divides a distribution into two parts of equal area, or the 50th percentile, that is, half the measurements are on one side of the median and half are on the other side when all measurements are arranged in order of magnitude. The median is unrelated to any arithmetic manipulation of values and is simply the middle one. For example, if there are 5 children whose ages are: 2, 2, 2, 2, 5, the median age is 2 years; this value is the middle one of all the values selected. The value of the statistic depends on the order or rank of the data and tells us more about the entire group than the mode does.

■ The mean, or arithmetic average, is the sum of the values divided by the total number of values. It is the most commonly used summary measure. Using the data on ages for the example of a median, sum all of the ages, then divide by five (the number of cases in the sample). The mean age of this group of children is 2.6 years. The difference between the mean and mode (2 versus 2.6) is that the mean is more sensitive to extreme numbers (the 5 year old) and reflects the value of every member in the group.

Since each measure of central tendency provides different information you can specify all three averages for a comprehensive description of a group.

Measures of Dispersion (Variability)

It is also useful to calculate measures of dispersion, or "variability," which indicate the level of spread of a group. Two such measures which may be of use to WIC administrators are:

■ Range. The range for numerical data is the difference between the largest and smallest observed values + 1 (range = highest - lowest + 1). The range for categorical data is the number of categories in the continuum, for example from "very satisfied" to "very dissatisfied."

■ **Variance.** The variance describes how close to a central measure, or how far away, the values of a distribution are. The standard deviation is often used to identify how elements of a group differ from the group mean. The measure of central tendency is more useful if the variance is small; that is, the values are not spread out.

Frequencies and Cross-Tabulations

These are other measures that can be easily calculated. A frequency is a simple count of the number of times something occurs. A cross-tabulation is a table of counts arranged into distributions. Note the accompanying example of a cross-tab of a WIC agency's caseload of pregnant women by age and race/ethnicity.